

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images,
Please do not report the images to the
Image Problem Mailbox.

THIS PAGE BLANK (USPTO)

~~TI - Relay apparatus e.g. router used in computer network, stores header information in table which is referred during packet routing for communication quality, forwarding priority and suitability~~

~~AB - A search table stores forwarding priority, quality and suitability corresponding to the packet forwarding destination information. When a packet is received by the network interface (30), the packet is stored in a buffer (12) and header information is stored in RAM (11). The packet forwarding destination is read from the header and the table is referred to determine the transmission parameters. An Independent claim is also included for the packet relay procedure in computer network.~~

~~PN - FR2794319 A 20001201~~

~~AP - FR20000002319 20000224~~

~~OPD - 1999-02-24~~

~~PR - JP19990046837 19990224; JP19990046579 19990224; JP19990046422 19990224; JP19990045959 19990224~~

~~PA - HITACHI LTD (JP)~~

~~IN - SUGAI KAZUO; YAZAKI TAKEKI; AIMOTO TAKESHI; SAKO YOSHIHITO; TANABE TOMOHIKO; MATSUYAMA NOBUHITO~~

~~EC - H04L12/56E~~

~~IC - H04L12/46 ; H04L12/56 ; H04L12/22~~

~~ICO - T04Q11/04S2B2G2 , T04Q11/04S2B7E~~

TI - Relay apparatus e.g. router used in computer network, stores header information in table which is referred during packet routing for communication quality, forwarding priority and suitability

AB - JP2000244574 NOVELTY - A search table stores forwarding priority, quality and suitability corresponding to the packet forwarding destination information. When a packet is received by the network interface (30), the packet is stored in a buffer (12) and header information is stored in RAM (11). The packet forwarding destination is read from the header and the table is referred to determine the transmission parameters.

- DETAILED DESCRIPTION - An INDEPENDENT CLAIM is also included for the packet relay procedure in computer network.
- USE - For e.g. router, bridge used in computer network for high speed packet transmission.
- ADVANTAGE - Highly reliable, secured packet transmission is performed, as the table that stores communication parameter corresponding to the destination is referred.

THIS PAGE BLANK (USPTO)

- DESCRIPTION OF DRAWING(S) - The figure shows the model of the network relay apparatus.

- RAM 11

- Buffer 12

- Network interface 30

- (Dwg.2/13)

PN - FR2794319 A1 20001201 DW200066 H04L12/46 000pp

- JP2000244574 A 20000908 DW200058 H04L12/56 014pp

OPD - 1999-02-24

PR - JP19990046837 19990224;JP19990045959

19990224;JP19990046422 19990224;JP19990046579 19990224

PA - (HITA-N) HITACHI INFORMATION TECHNOLOGY KK

- (HITA) HITACHI LTD

- (HITA-N) HITACHI COMPUTER ELECTRONICS KK

IN - AIMOTO T; MATSUYAMA N; SAKO Y; SUGAI K; TANABE T; YAZAKI
T

IC - H04L12/22 ;H04L12/46 ;H04L12/56 ;H04L29/06

AN - 2000-607336 [58]

THIS PAGE BLANK (USPTO)

【特許請求の範囲】

【請求項1】 ネットワークと接続されたネットワークインタフェース部と、

入力パケットを記憶するパケットバッファと、パケットの転送元又は転送先を含む情報に対応するアクション情報をエントリとして入力又は出力回線番号毎に分割して設定されたフロー検索テーブルとを有し、前記ネットワークインタフェース部から入力されたパケットのルーティングを行うルーティング処理部と、

装置内部を管理するルーティング管理部と、

前記ルーティング管理部及び複数の前記ルーティング処理部の各々を接続する接続部とを備え、

前記ネットワークインタフェース部は、ネットワークから入力された入力パケットを前記ルーティング処理部へ出力し、

前記ルーティング処理部は、

前記ネットワークインタフェース部からの入力パケットをバッファメモリに記憶し、

記憶されたヘッダ情報に基づいて、パケットバッファに記憶された入力パケットの転送先を検索し、

前記フロー検索テーブルを参照して、パケットの入力又は出力回線番号に対応するエントリのみを検索して読み出し、

パケットの転送元又は転送先を含む情報と読み出されたエントリ内の比較条件との一致を判定し、

両者が一致すると判定された場合に、エントリ内のアクション情報により、パケットの転送の優先度又は転送の可否等の通信品質に関する制御を決定し、

前記パケットバッファに記憶された入力パケット及び前記ヘッダ情報により作成された出力パケットを、前記接続部又は前記ネットワークインタフェース部に出力するようにしたネットワーク中継装置。

【請求項2】 前記ルーティング処理部は、

前記フロー検索テーブル内のエントリを指定するエントリポインタを前記アドレス情報毎に分割して設定したエントリポインタテーブルをさらに備え、

パケット内のアドレス情報に対応するエントリポインタのみを読み出し、

読み出されたエントリポインタが指定するエントリを前記フロー検索テーブルから読み出し、

パケット内のアドレス情報と、読み出されたエントリ内の情報との一致を判定し、

一致と判定された場合に、エントリ内のアクション情報により、パケットの転送の優先度又は転送の可否等の通信品質に関する制御を行うことを特徴とする請求項1に記載のネットワーク中継装置。

【請求項3】 前記アクション情報は、通過/廃棄等の処理を行うフィルタ制御、カプセル化する又はしない等の処理を行うトンネル制御、及び、遅延クラス若しくは廃棄クラスによる優先処理又は帯域に関する処理を行うQ

oS制御のいずれかひとつ又は複数の制御に対応することを特徴とする請求項1又は2に記載のネットワーク中継装置。

【請求項4】 前記パケットバッファと独立して高速読出し及び書込み可能であり、入力パケットに関するヘッダ情報を記憶するヘッダメモリをさらに備えた請求項1乃至3のいずれかに記載のネットワーク中継装置。

【請求項5】 ネットワークと接続されたネットワークインタフェース部と、前記ネットワークインタフェース部から入力されたパケットのルーティングを行うルーティング処理部と、装置内部を管理するルーティング管理部と、前記ルーティング管理部及び複数の前記ルーティング処理部の各々を接続する接続部とを備えたネットワーク中継装置において、ネットワークから入力された入力パケットを転送先に出力するネットワーク中継方法であって、

前記ルーティング処理部は、

フロー検索テーブルに、パケットの転送元又は転送先を含む情報に対応するアクション情報をエントリとして入力又は出力回線番号毎に分割して設定し、

入力パケットをバッファメモリに記憶し、

記憶されたヘッダ情報に基づいて、パケットバッファに記憶された入力パケットの転送先を検索し、

前記フロー検索テーブルを参照して、パケットの入力又は出力回線番号に対応するエントリのみを検索して読み出し、

パケットの転送元又は転送先を含む情報と読み出されたエントリ内の比較条件との一致を判定し、両者が一致すると判定された場合に、エントリ内のアクション情報により、パ

ケットの転送の優先度又は転送の可否等の通信品質に関する制御を決定し、

前記パケットバッファに記憶された入力パケット、及び、前記ヘッダ情報により作成された出力パケットを、前記接続部又は前記ネットワークインタフェース部に出力するようにしたネットワーク中継方法。

【請求項6】 前記ルーティング処理部は、

さらに、エントリポインタテーブルに、前記フロー検索テーブル内のエントリを指定するエントリポインタを前記アドレス情報毎に分割して設定し、

パケット内のアドレス情報に対応するエントリポインタのみを読み出し、

読み出されたエントリポインタが指定するエントリを前記フロー検索テーブルから読み出し、

パケット内のアドレス情報と、読み出されたエントリ内の情報との一致を判定し、

一致と判定された場合に、エントリ内のアクション情報により、パケットの転送の優先度又は転送の可否等の通信品質に関する制御を行うことを特徴とする請求項5に記載のネットワーク中継方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、ネットワーク中継装置及びネットワーク中継方法に係る。本発明は、特に、コンピュータネットワークシステムにおけるルータ等のネットワーク中継装置において、入力されたパケットの転送先を高速に検索することができるネットワーク中継装置及びネットワーク中継検索方法に関する。

【0002】

【従来の技術】一般に、ネットワークシステムにおいては、複数のネットワーク間を接続するためにルータやブリッジ等のネットワーク中継装置が用いられる。ルータは、接続されているネットワークやサブネット等から受け取ったパケットの宛先アドレスを調べてパケットの転送先を決定し、転送先のルータ又はホストが接続されたネットワークやサブネット等にパケットを転送する。

【0003】図13に、従来のネットワーク中継装置の構成図を示す。この図において、ルータ100は、ルーティング制御部(RM)110、ルータバス120、ネットワークインタフェース部(NIF)130、ポート140等を備える。各ポート140は、適宜のネットワーク150に接続される。

【0004】ネットワークインタフェース部130は、ポート140に接続されたネットワークからパケットを受け取り、受け取ったパケットをルータバス120経由でルーティング制御部110に送信する。ルーティング制御部110は、ルーティング情報を保持するルーティングテーブルを備え、このルーティング情報を用いて受け取ったパケットの宛先から転送先のネットワーク150を決定し、そのネットワーク150が接続されるポート140のネットワークインタフェース部130にパケットを送信する。ルーティング制御部110からパケットを受け取ったネットワークインタフェース部130は、そのパケットを転送先のネットワーク150に送出する。なお、ルーティング制御部110は、受け取ったパケットのヘッダ情報に基づいてルーティングテーブルに保持するルーティング情報を更新・保守するとともに、ルータ100全体の管理機能を備えている。

【0005】ここで、ルータがパケットを受け取ったときに、次に転送するアドレス及びパケットを出力するポートを検索する経路検索処理について説明する。通常、経路検索には、構成定義情報及びルータ間での送受により得られた情報等から作成される経路検索テーブル(ルーティングテーブル)が用いられる。ルーティングテーブルは、ネットワークアドレスとネットワークマスク長等の組を検索のキーとして、出力ポート、次ホップアドレス、及び、ネットワークが直接接続されているか否かの情報(次ホップ情報)等を検索するためのテーブルである。

【0006】また、他の従来技術として、特開平05-

199230号公報(米国特許第5,434,863号参照)には、ルーティング処理の高速性を損なうことなく、ネットワークの規模に柔軟に対応できるようにしたインターネットワーク装置及び通信ネットワークシステムが記載されている。この装置及びシステムでは、ルータ管理部と、ルーティングを行う複数のルーティングアクセラレータモジュールとが、高速バスで結合される構成となっている。また、各々のルーティングアクセラレータには、複数の通信ポートが互いに独立に接続されている。このような従来技術によると、複数のルーティングアクセラレータにより、高速ルーティングを可能とし、また、ルーティングアクセラレータを増設すれば、小規模から大規模ネットワークへの対応が容易に実現できる。

【0007】

【発明が解決しようとする課題】しかしながら、従来のルータでは、近年出現してきたの高速LAN(Local Area Network)や、広帯域ISDN及びATM(Asynchronous Transfer Mode)など高速回線に対応するのは困難である。また、従来のルータでは、ルーティングを行う手段が一ヶ所であるためルーティング処理がネックとなり、サポートできるポート数や通信トラフィックに限界がある。従って、ルータのポートメニュー等の構成を小規模から大規模までスムーズに拡張すること、ポートのトラフィックや数に応じて性能を向上させることが困難である。

【0008】また、インターネットユーザの増加に伴い、ルータが検出しなければならないフローの数が増加している。そのため、ルータでは多数のフロー条件を設定できなくてはならない。また、インターネットを流れるトラフィックの急増、回線速度の高速化により、ルータにおける1パケットあたりの処理時間の短縮が必要となる。そのため、ルータは、設定されるフロー条件数が増加した場合にも、QoS(Quality of Service、サービス品質)制御やフィルタリングを高速に行えなければならない。

【0009】一方、フロー条件の設定に関しては、ルータの管理者が望む多種多様なフロー条件に柔軟に対応できることが望ましい、従来の技術では、かかる観点は検討されていない。

【0010】本発明は、以上の点に鑑み、高通信品質(QoS)、高信頼性、セキュリティを保証して高速にルーティングするネットワーク中継装置及び方法を提供することを目的とする。

【0011】また、本発明は、ユーザを識別する情報やプロトコル情報や優先度情報等のフロー条件を大量に設定でき、回線速度の高速化やフロー条件の増加に対応し、高速にフロー検出し、高速にQoS制御やフィルタリング等の通信品質に関する制御を実現することを目的とする。また、フロー条件の記述性を向上させ、転送元

及び転送先の情報等の組合せ条件により、優先制御、廃棄制御、帯域制御等の多種多様なフロー条件に柔軟に対応した制御を高速に処理することを目的とする。

【0012】

【課題を解決するための手段】本発明の第1の解決手段によると、ネットワークと接続されたネットワークインタフェース部と、入力パケットを記憶するパケットバッファと、パケットの転送元又は転送先を含む情報に対応するアクション情報をエントリとして入力又は出力回線番号毎に分割して設定されたフロー検索テーブルとを有し、前記ネットワークインタフェース部から入力されたパケットのルーティングを行うルーティング処理部と、装置内部を管理するルーティング管理部と、前記ルーティング管理部及び複数の前記ルーティング処理部の各々を接続する接続部とを備え、前記ネットワークインタフェース部は、ネットワークから入力された入力パケットを前記ルーティング処理部へ出力し、前記ルーティング処理部は、前記ネットワークインタフェース部からの入力パケットをバッファメモリに記憶し、記憶されたヘッダ情報に基づいて、パケットバッファに記憶された入力パケットの転送先を検索し、前記フロー検索テーブルを参照して、パケットの入力又は出力回線番号に対応するエントリのみを検索して読み出し、パケットの転送元又は転送先を含む情報と読み出されたエントリ内の比較条件との一致を判定し、両者が一致すると判定された場合に、エントリ内のアクション情報により、パケットの転送の優先度又は転送の可否等の通信品質に関する制御を決定し、前記パケットバッファに記憶された入力パケット及び前記ヘッダ情報により作成された出力パケットを、前記接続部又は前記ネットワークインタフェース部に出力するようにしたネットワーク中継装置を提供する。

【0013】本発明の第2の解決手段によると、ネットワークと接続されたネットワークインタフェース部と、前記ネットワークインタフェース部から入力されたパケットのルーティングを行うルーティング処理部と、装置内部を管理するルーティング管理部と、前記ルーティング管理部及び複数の前記ルーティング処理部の各々を接続する接続部とを備えたネットワーク中継装置において、ネットワークから入力された入力パケットを転送先に出力するネットワーク中継方法であって、前記ルーティング処理部は、フロー検索テーブルに、パケットの転送元又は転送先を含む情報に対応するアクション情報をエントリとして入力又は出力回線番号毎に分割して設定し、入力パケットをバッファメモリに記憶し、記憶されたヘッダ情報に基づいて、パケットバッファに記憶された入力パケットの転送先を検索し、前記フロー検索テーブルを参照して、パケットの入力又は出力回線番号に対応するエントリのみを検索して読み出し、パケットの転送元又は転送先を含む情報と読み出されたエントリ内の

比較条件との一致を判定し、両者が一致すると判定された場合に、エントリ内のアクション情報により、パケットの転送の優先度又は転送の可否等の通信品質に関する制御を決定し、前記パケットバッファに記憶された入力パケット、及び、前記ヘッダ情報により作成された出力パケットを、前記接続部又は前記ネットワークインタフェース部に出力するようにしたネットワーク中継装置を提供する。

【0014】

【発明の実施の形態】図1に、本発明に係るネットワーク中継装置の構成図を示す。この図において、ルータ1は、ルーティング処理部(RP, Routing Processor)10、スイッチ部(CS, Crossbar Switch)20、ネットワークインタフェース部(NIF, Network Interface)30、ポート40、ルーティング管理部(RM, Routing Manager)60、電源(PS, Power Supply)70等を備える。各ポート40は、適宜のネットワーク50に接続される。ネットワーク50としては、例えば、イーサネット等のLAN、WAN、ATM等がある。なお、装置の信頼性向上させるため、電源70及び共通部等を適宜二重化することができる。

【0015】ルーティング制御のための機能は、ルーティング機能を実行するルーティング処理部10及びルータ1の管理を行うルーティング管理部60に分かれる。さらに、ルータ1は、ひとつ又は複数のネットワークインタフェース部30を有するルーティング処理部10を、複数備えている。ルーティング管理部60は、ルータ1全体の管理機能を備えるとともに、ルート計算処理を実行し、また、他のルータとの間でルーティング情報を送受するとともに、ルータ内の各ルーティング処理部10にルーティング情報を配布する。ルーティング管理部60は、ここでは、二重化構成が採用される。スイッチ部20は、クロスバスイッチ等のスイッチを備え、ルーティング処理部10相互間又はルーティング処理部10とルーティング管理部60との間の通信及び交換を行う。スイッチ部20は、ここでは、二重化構成が採用される。なお、スイッチ部20に代えてバス等により接続しても良い。また、クロスバスイッチを用いた場合、ルーティング管理部60及び複数のルーティング処理部10のうち1組により接続経路が占有されることなく、同時に複数の組が送受信を行うことができる。

【0016】ルーティング処理部10は、接続されたネットワークインタフェース部30によりパケット転送を行う。ルーティング処理部10は、また、他のルーティング処理部10に接続されたネットワーク50にパケットを転送する場合は、スイッチ部20を介して該当するルーティング処理部10にパケットを転送する。ルーティング処理部10は、各機能がハードウェア構成により、高速動作をするよう設計される。ルーティング処理部10は、さらに詳細には、例えば、スイッチング処理、ル

ート検索、フォワーディング、フィルタリング、QoS、IPマルチキャスト等の機能を有する。ルーティング処理部10は、自己のルーティング処理部10内のネットワークインタフェース部30の各ポート40、他の各ルーティング処理部10及びルーティング管理部60等の各々について、適宜入力バッファ及び出力バッファを備える。ネットワークインタフェース部30は、ひとつ又は複数のポート40を有し、ネットワーク50とルーティング処理部10とのインタフェースを制御する。

【0017】図2に、ルーティング処理部の内部構成を表したネットワーク中継装置の動作説明図を示す。この図を用いて、経路検索と、経路検索の結果得られた宛先にパケットを転送する動作について、ルーティング処理部10の内部構成図を参照して説明する。ルーティング処理部10は、転送エンジン13、検索エンジン14、ヘッダRAM11、パケットバッファ12、経路テーブル15、ARPテーブル（アドレス検索テーブル）（ARP, Address Resolution Protocol）16、フィルタ/QoSテーブル（フロー検索テーブル）17を備える。転送エンジン13は、例えば、パケットの入出力処理を行う。検索エンジン14は、主に、パケットのヘッダ情報に基づき経路検索処理及びQoS制御等のフロー検索処理を行う。検索エンジン14は、専用LSI等により、高速処理が可能なハードウェアで構成される。

【0018】パケットバッファ12は、転送エンジン13がルーティング処理部10に入力されたパケットを転送するまでの間、入力されたパケットを格納しておく。ヘッダRAM11は、入力されたパケットのヘッダのみを抽出して記憶する。ヘッダRAM11は、読み出し/書込速度の高速なメモリを用いる。この実施の形態では、ネットワークから受信したパケットや他のデータ処理装置から転送されたパケットを格納するバッファメモリとは別に、パケットバッファ12と非同期にアクセス可能なヘッダRAM11を設け、パケットをパケットバッファ12へ格納すると同時に該パケットのヘッダ部をヘッダRAM11へも格納（コピー）するようにした。転送エンジン13及び検索エンジン14等の各々のプロセッサは、ヘッダRAM11を使用してパケットのヘッダ部を取り込み、ヘッダ部を解析している間、パケットバッファ12に対するパケットの読み書きが可能になり、あるパケットのヘッダ解析処理と他のパケットの転送処理等の並列動作が達成される。

【0019】したがって、検索エンジン14のプロセッサがヘッダ情報をヘッダRAM11から読み出している間、パケットバッファ12はプロセッサが使用しないので、転送エンジン13によるパケット送信又は転送のためのアクセスが可能になり、検索エンジン14と転送エンジン13とのパケットバッファ12へのアクセス競合が起こらないようにすることができる。なお、ネットワーク側からの受信パケット及びそのヘッダ部を格納する

領域と、スイッチ部20側からの転送パケット及びそのヘッダ部を格納する領域とを、それぞれ別構成としてもよい。このように別構成にすると、パケットの管理が容易になる。

【0020】また、経路テーブル15、ARPテーブル16及びフィルタ/QoSテーブル17は、それぞれ独立に分割して構成される。これにより、検索エンジン14が各テーブルを個別にアクセスして読み出し又は書込を行うことで、ルーティング情報、QoS等を高速に検索することができる。さらに、ルーティング処理の高速化の実現のために、パイプライン処理を実行することができる。各テーブルの詳細及びパイプライン処理については、後述する。

【0021】図3に、ネットワーク中継装置の動作概要のシーケンス図を示す。まず、ネットワークからボードを介して、第1のネットワークインタフェース部30にパケットが入力されると、第1のネットワークインタフェース部30は、これを転送エンジン13に送信する。転送エンジン13は、受信したパケットをパケットバッファ12に記憶する。また、転送エンジン13は、入力パケットのヘッダのみを抽出し、内部ヘッダを付加してヘッダ情報を形成し、ヘッダ情報をヘッダRAM11に記憶する。内部ヘッダについては、後述する。

【0022】検索エンジン14は、ヘッダRAM11にアクセスしてヘッダ情報を読み出す。なお、転送エンジン13が、ヘッダRAM11に記憶されたヘッダ情報を検索エンジン14に送信するようにしてもよい。検索エンジン14では、ヘッダ情報により、宛先のルータ・RIP・ポートの各番号又はアドレス、MACアドレス等の次の転送経路に関する情報、及び、QoS制御情報等の通信品質に関する制御についての情報等を適宜検索する。検索エンジン14は、検索された各番号・アドレス情報等の転送先情報及びQoS情報等のアクション情報を含む転送制御情報をヘッダRAM11に書き込む。なお、検索エンジン14は、転送エンジン13に転送制御情報を送信するようにしてもよい。

【0023】転送エンジン13では、パケットバッファ12に記憶されたパケット及びヘッダRAM11に記憶されたヘッダ情報（転送制御情報を含む）に基づいて、出力パケットを作成する。そして、転送エンジン13は、作成された出力パケットを転送先に出力する。この際、転送エンジン13は、転送経路が他のルーティング処理部10に属するものである場合は、該当する他のルーティング処理部10へのバッファにキューイングし、自己のルーティング処理部10のネットワークインタフェース部30に属するものである場合は、対応するポート40のバッファにキューイングする。

【0024】なお、ルーティング処理部10により検索される転送経路は、ひとつに限らず、複数経路として各経路に同報転送することもできる。この場合、複数の各

経路毎に適宜のバッファにキューイングされる。

【0025】以下に、ルーティング処理部の詳細な構成及び動作を説明する。まず、各メモリの説明をする。図4に、パケットバッファ12及びヘッダRAM11の説明図を示す。

【0026】図4(A)に、パケットバッファ12に記憶されるパケットのフォーマットの一例を示す。パケットバッファ12には、ネットワーク50又はスイッチ部20等から、パケットが入力される。このパケットのフォーマットは、例えば、IPパケットに、レイヤ2・MACヘッダ401が付加されたものである。IPパケットは、例えば、レイヤ3・IPヘッダ402、レイヤ4・ヘッダ403及びペイロード404を含む。

【0027】レイヤ2・MACヘッダ401は、例えば、パケットを直前に送出したルータの物理アドレス(ハードウェアアドレス)である送信元MACアドレス(Source Address Media Access Control, SAMAC)と、パケットを次に受信するルータの物理アドレスである宛先MACアドレス(Destination Address Media Access Control, DAMAC)等を有する。レイヤ3・IPヘッダ402は、送信元アドレス(送信端末のアドレス)である送信元IPアドレス(Source IP Address: 以下「SI P」という。)と、宛先アドレス(受信端末のアドレス)である宛先IPアドレス(Destination IP Address: 以下「DIP」という。)等を有する。レイヤ4ヘッダ403は、プロトコル(=上位アプリケーション)を表す送信元ポート(Source Port: 以下「SPORT」という。)と宛先ポート(Destination Port: 以下「DPOR T」という。)等を有する。ペイロード404は、ユーザデータを含む。各ヘッダとしては、上述の他に優先度を表すTOS(Type of Service)、IPプロトコルの上位プロトコル等の情報も格納される場合があるが、上述した各情報と同様に処理することができる。

【0028】また、図4(B)に、ヘッダRAMに記憶されるヘッダ情報のフォーマット一例を示す。ヘッダ情報は、例えば、パケットフォーマット中のレイヤ2・MACヘッダ401及びレイヤ3・IPヘッダ402に、制御情報として内部ヘッダ405を付加したものである。この内部ヘッダ405は、例えば、入力回線番号、出力回線番号、QoS制御情報等を有する。なお、ルータ内部での内部パケットフォーマットは、ネットワークのパケットのフォーマットに内部ヘッダ405が付加される。その際、パケットバッファ12に記憶された情報とヘッダRAM11に記憶された情報とにより内部パケットが形成されることができる。また、パケットバッファ12に、内部ヘッダ405を含めた内部パケットフォーマットの形式で記憶するようにして、パケットバッファ12のみの情報から内部パケットを転送するようにしても良い。また、検索エンジン14により検索された転送先情報及びアクション情報等の転送制御情報は、この

内部ヘッダ405に書き込まれることができる。

【0029】つぎに、図5に、経路検索に用いられる各テーブルの説明図を示す。図5(A)に示すように、経路テーブル15の各エントリは、例えば、宛先のIPアドレス501、次ルータのIPアドレス502、自ルータの送出RP番号503及び送出ポート番号504等を含む。また、図5(B)に示すように、ARPテーブル16の各エントリは、例えば、次ルータのIPアドレス502、次ルータのMACアドレス506等を含む。また、図5(C)に示すように、フィルタ/QoSテーブル17の各エントリは、例えば、IPヘッダ/レイヤ4ヘッダの値(範囲)507及びアクション508等を含む。ここで、アクション508としては、通過又は廃棄をする処理であるフィルタ、カプセル化する又はカプセル化しない処理であるトンネル、QoS等がある。特にQoSについては、後述する。

【0030】図6に、ルーティング処理部による高速化処理の説明図を示す。この図を参照して、ギガビット等の高速な回線速度に追従できるようにしたパケット転送の実現化方法について説明する。ここでは、ルーティング処理の平行化/パイプライン化により、高速化を実現する。以下に、図4及び図5に示されるフォーマットを参照して、動作を説明する。

【0031】ルーティング処理は、主に、受信処理①、入力検索処理②、出力検索処理③、送信処理④に分けられる。まず、受信処理①では、転送エンジン13は、ネットワークインタフェース部30からの受信パケットを受け取る。パケットバッファ12には、前述のとおり入力パケット又は内部ヘッダを付加した内部パケットフォーマットのパケットが記憶される。また、入力パケットのレイヤ2・MACヘッダ401及びレイヤ3・IPヘッダ402に内部ヘッダ405を付加してヘッダ情報とし、ヘッダ情報がヘッダRAM11に記憶される。ヘッダRAM11は、パケットバッファ12と独立に高速に読み出し及び書き込みが可能であり、ヘッダ情報のみを記憶することで、記憶容量を削減し、一層高速に処理をすることができる。検索エンジン14は、抽出されたヘッダ情報を、適宜のタイミングでアクセスすることができる。

【0032】つぎに、入力検索処理②では、検索エンジン14は、ヘッダ情報からレイヤ3・IPヘッダ402中にある宛先IPアドレスを抽出し、このアドレスに基づき経路テーブル15を参照して、次ルータのIPアドレス502、自ルータの送出RP番号503及び送出ポート番号504等を検索する。さらに、検索エンジン14は、受信したヘッダ情報からレイヤ3・IPヘッダ402及びレイヤ4・ヘッダ403等の比較情報に基づき、フィルタ/QoSテーブル17を参照して、QoS等の入力側の各種アクション508を検索する。これらの入力側フィルタ/QoS検索及び経路検索は、各々の

テーブルが別個に設けられているため、並列して実行することができる。

【0033】つぎに、出力検索処理③では、検索エンジン14は、入力検索処理②で求められた次ルータのIPアドレスを抽出し、このアドレスに基づきARPテーブル16を参照して、次ルータのMACアドレス506等を検索するとともに、フィルタ/QoSテーブル17を参照して、QoS等の出力側の各種アクション508を検索する。これらの出力側フィルタ/QoS検索及び回線表/ARP検索は、各々のテーブルが別個に設けられているため、並列して実行することができる。求められた次の転送先に関する番号・アドレス情報等の転送先情報及びQoS制御情報等のアクション情報を含む転送制御情報は、ヘッダRAM11に記憶される。転送制御情報は、例えば、ヘッダ情報中の内部ヘッダ405又はその他の位置に適宜書き込むことができる。

【0034】つぎに、送信処理④では、出力検索処理③が検索した転送制御情報を含むヘッダ情報をヘッダRAM11から読み出し、ヘッダ情報とパケットバッファ12に基づき出力パケットを作成して、ネットワークインタフェース部30又は他のルーティング処理部10若しくはルーティング管理部60に対するバッファにキューイングする。

【0035】図7に、ハードウェアによる検索エンジンの構成図の一例を示す。検索エンジン14は、経路テーブル15、ARPテーブル16、フィルタ/QoSテーブル17等の各テーブルについて、例えば、木構造検索により必要なデータを検索することができる。ここでは、ハードウェアで構成した検索エンジン14の処理部の一例として、経路テーブル15を用いて転送先経路を検索するための経路検索処理部について説明する。

【0036】経路検索処理部213は、木構造検索回路2130と、読み込みアドレス生成回路2131と、経路検索処理制御回路2132を備える。木構造検索回路2130は、経路テーブル15等の各テーブルに格納された2のp乗分木構造を検索し、次に読み込むべきノードのポインタの生成、受信パケットの宛先IPアドレスの検査ビット値の抽出、木構造検索の終了判定、検索結果である経路情報の候補の更新等を行う。また、読み込みアドレス生成回路2131は、木構造検索回路2130から出力される読み込むべきノードへのポインタ、および検査ビット値に従い、実際に読み込むノードの一部のワードのメモリアドレスを生成する。また、経路検索処理制御回路2132は、経路検索処理部213全体の制御(各回路の動作タイミングおよび動作状態管理など)を行う。

【0037】次に、経路検索処理部213の動作について説明する。木構造検索回路2130は、ヘッダRAM11から受信パケットの宛先IPアドレスを受け取り、この宛先IPアドレスとノードのマスク長の値から次ノ

ードへのポインタを生成して、読み込みアドレス生成回路2131に渡す。また、木構造検索回路2130は、ノードのマスク長で示される宛先IPアドレスの検査ビット位置の値(検査ビット値)を抽出して、読み込みアドレス生成回路2131に渡す。

【0038】読み込みアドレス生成回路2131はこのノードへのポインタと、検査ビット値と、経路検索処理制御回路2132からのタイミング信号を用いて、読み出すべきノードデータが格納されているメモリアドレスを生成し、メモリ制御回路2132へ送信し、メモリ制御回路2132はこのメモリアドレスと経路検索処理制御回路2132からのタイミング信号を用いてメモリ制御信号を生成し、経路テーブル15へ転送する。上記のメモリ制御信号を受信した経路テーブル15は、対応するノードデータを信号線215を用いて木構造検索回路へ転送する。

【0039】木構造検索回路2130はこのノードデータを用いて、検索し、木構造検索を終了すると判定した場合は、木構造検索終了信号を経路検索処理制御回路2132へ出力し、経路検索処理制御回路2132は、木構造検索回路2130内に保持された経路情報のうちのエントリ有りフラグを調べ、その値が0の場合は経路検索処理を終了し、転送エンジン13等へ検索結果無しという通知をする。エントリ有りフラグの値が1の場合は、経路情報を出力し検索処理を終了し、次のパケット処理の制御を行う。

【0040】つぎに、図8に、パイプライン制御による高速処理の説明図を示す。図示されるように、受信処理①、入力検索②、出力検索処理③、転送処理④の各処理をパイプラインで処理し、各処理部が常に動作するように制御することにより、ルーティング処理を高速化することができる。ここでは、さらに、入力検索②において、入力フィルタ処理(入力フィルタ/QoS検索)と、経路表検索(経路検索)を平行して実行している。また、出力検索処理③、出力フィルタ処理(出力フィルタ/QoS検索)と、出力回線表検索(出力回線表/ARP検索)を平行して実行している。パイプライン構成は、この図に限らず、適宜の順序で実行することができる。

【0041】パイプライン処理では上述の各処理部のうちの処理部1がエントリNの処理1を終了すると、処理1より後の処理2を行う処理部2がエントリNの処理2を終了したか否かに拘わらず、処理部1がエントリN+1の処理を開始する。この様なパイプライン処理を行うことにより、エントリNを処理する時間は、1つの処理の時間となり、処理速度は4倍となる。前述の例では、フロー検索が4つの処理に分割されパイプライン処理したが、P個の処理に分割し、パイプライン処理できればP倍の性能となる。

【0042】つぎに、図9に、フロー検索処理の説明図

を示す。一般に、ルータ等のネットワーク中継装置は、予めコネクションを設定していないので、ATM交換機の場合のコネクション情報テーブルやコネクション情報テーブル内のQoS制御情報を持っていない(パケット型通信)。このため、ルータでQoS制御を行うためには、ATM交換機と同様の優先転送機能の他に、入力パケット毎にヘッダ内の情報等によりQoS制御情報を検索するフロー検索手段が必要となる。ここでは、一例として、フロー検索手段により、検索されたQoS制御情報に対し優先転送機能を適用する。なお、ここでは、ヘッダ内の情報等の情報を組み合わせて作成したパケット識別の条件をフロー条件と呼び、フロー条件に一致する一連のトラヒックをフローと呼び、また、フロー条件に入力パケットが一致するか否かを判定してQoS制御情報や転送可否情報等のアクション情報を検出することをフロー検索と呼ぶ。

【0043】本実施の形態では、QoS制御を入力側ルーティング処理部10-1及び出力側ルーティング処理部10-2に、それぞれ設けると共に、スイッチ部20にQoS機能を持たせるようにした。入力ルーティング処理部10-1では、入力検索フローとしてフィルタフロー検索911、トンネルフロー検索912、QoSフロー検索913を備える。同様に、出力ルーティング処理部10-2では、出力検索フローとしてフィルタフロー検索921、トンネルフロー検索922、QoSフロー検索923を備える。スイッチ部20は、優先度により送信する順序を選択するアービトレーション機能を備えることで、QoS機能を持たせるようにした。なお、スイッチ部20にも、ルーティング処理部10-1及び10-2と同様に、フィルタフロー検索、トンネルフロー検索、QoSフロー検索を備えることもできる。

【0044】フィルタフロー検索911及び921では、パケットを通過させるか又は廃棄するかを求める。トンネルフロー検索912及び922では、パケットをカプセル化するか否かを求め、カプセル化する場合は、カプセル化のソフトウェア処理を実行する。

【0045】QoSフロー検索913及び923では、例えば、パケットの優先制御、廃棄制御、帯域制御等がある。優先制御とは、重要度の高いデータやリアルタイム系データ等を優先的に送信する制御である。廃棄制御とは、トラフィックが多い場合・障害発生等において、重要度の低いデータを廃棄することにより、重要なデータの消失を防止する制御である。また、帯域制御とは、回線内を複数の帯域に分割したり、帯域幅を変更する制御である。たとえば、優先クラスと廃棄クラスのマトリクスを用いてトラフィック制御することで、優先及び廃棄制御を行うことができる。この場合、優先クラスにより、例えば、HNA/SNA、音声、動画等は遅延小とし、FTP、メール、WWWウェブ等は、遅延大とする制御を行うことができる。また、廃棄クラスにより、例

えば、制御パケットは廃棄率を小さくし、音声、動画等は、廃棄率を大とするように制御することができる。

【0046】つぎに、スイッチ部20でのQoS制御について説明する。ルーティング処理部10から送られてきたパケットには、制御情報中にQoS制御情報が含まれている。スイッチ部20では、特に出力側において、このQoS情報による優先制御等を行う。実際には、例えば、優先順位毎にキューを備えて出力制御を行うこと等により制御が可能である。これにより、一段と高品質の通信及び転送を行うことができる。

【0047】図10に、フロー検索テーブルの説明図を示す。このフロー検索テーブルは、上述のフィルタ/QoSテーブル17に対応するものである。ここでは、一例として、図示のように、比較フィールド101として、転送元IPアドレス、宛先IPアドレス、パケット長、IP優先度、IP上位プロトコル、送達確認フラグ、転送先TCP/UDPポート、宛先TCP/UDPポートを含む。また、アクションフィールド102としては、フィルタ(通過/廃棄)、トンネル(カプセル化する/しない)、QoS(遅延クラス、廃棄クラス、帯域等)が記憶されている。

【0048】次に、QoSフロー検索の具体的方法を説明する。ここでは、一例としてQoSフロー検索に注目して説明するが、フィルタフロー検索又はトンネルフロー検索についても、同様に実施することができる。なお、アクションフィールド102には、これら各フローの制御情報が混在して記憶されていても良いし、また、各フロー毎にフロー検索テーブルが設けられるようにしても良い。

【0049】まず、リニアサーチ方式について説明する。アクションのひとつとしてのパケットのQoS制御情報を判定する際、予め設定されたエントリをエントリテーブルの上から順に読みだす。そして、パケットのヘッダ部等の値とエントリの比較フィールド101内の有効なフロー条件に全て一致したか否かを判定する。なお、一致した場合にはエントリ内のアクションフィールド102内のQoS制御情報をパケットのQoS制御情報と判定し、QoSフロー検索を終了する。フロー条件に一致する場合が検索されると、アクションフィールド102内のQoS制御情報をQoS制御情報と決定し次のエントリの検索を実行せずフロー検索を終了する。

【0050】以上のようなリニアサーチ方式では、大量のエントリが設定されるネットワークにおいては高速にQoS制御やフィルタリングを実行することが困難となる場合がある。そこで、本実施の形態のフロー検索方式では、大量のエントリが設定された場合でも、リニアサーチ方式と比較して高速にフロー検索を行うことができる、入力回線限定方式等を採用するとよい。以下、入力回線限定方式の概要を説明する。入力回線限定方式では、リニアサーチ方式の比較フィールドを構成する入

力回線番号に一致するエントリだけを検索し、高速化を図るようにする。

【0051】図11に、第1の入力回線限定方式の説明図を示す。第1の入力回線限定方式においては、リニアサーチ方式の比較フィールドから入力回線番号と入力回線番号有効ビットを削除したエントリ511-iを入力回線毎に設定しておく。フロー条件部521-iは、例えば、送信元あるいは宛先ユーザを識別する条件はSIPとDIPの上限値と下限値であるSIP上限値501、SIP下限値502、DIP上限値503、DIP下限値504と、SIPとDIPの上限値と下限値が有効であることを示すIP有効ビット562と、送信元ポートであるSPORT505と、宛先ポートであるDPORT506と、前記SPORT505とDPORT506が有効であることを示すポート有効ビット563等を含む。QoS制御情報部530-iは、例えば、優先転送機能で使用するQoS制御情報であるQoS制御情報507を含む。フロー条件である入力回線番号が一致したエントリ511-iだけ検索するためエントリ511-i内に入力回線番号は必要ない。フロー検索時には、パケットが入力した入力回線に割り当てられたエントリ511-iのみ検索する。

【0052】ここで、上述の第1の入力回線限定方式では、入力回線番号に関係ないエントリ511-iを設定する(例えば「全ての入力回線から入力されたTelenetのトラヒックは高優先」と設定する)場合、エントリ511-iを入力回線数(N)設定する必要がある。エントリテーブルを実現するメモリの効率が悪くなる場合がある。そこで、以下に、一層高速な入力回線限定方式について説明する。

【0053】図12に、第2の入力回線限定方式の説明図を示す。第2の入力回線限定方式では、エントリテーブル750のアドレスであるリスト540を、入力回線毎にリストテーブル760に設定しておく。例えば、リストテーブルアドレスが“1”のリスト540-11はエントリ511-1のアドレスであり、リストテーブルアドレスが“2”のリスト540-12はエントリ511-Hのアドレスである。フロー検索時には、パケットが入力された入力回線に割り当てられたリスト540だけ読みだし、このリスト540がポイントするエントリ511-iを読み出す。ビット幅の小さなリスト540(例えば、1024エントリ持った時にも10bit程度)を入力回線毎に所持し、ビット幅の大きなエントリ511-iを各入力回線で共有すれば、エントリテーブルを実現するメモリを有効に使用することができる。このため、高速化を実現しつつ、多数のエントリ511-iを設定することが可能となる。

【0054】フロー検出方式の他の実施例として出力回線限定方式がある。出力回線限定方式では、フロー条件である出力回線番号が一致するエントリ511-iのみを、上述の入力限定方式と同様に処理することで、フ

ロー検出の高速化を実現する。また、フロー条件としてヘッダ情報内の入力回線番号の代わりにSAMACを用いた際のSAMAC限定方式がある。SAMAC限定方式では、SAMACのグループであるSAMACグループを定義し、SAMACグループの識別子であるSAMAC識別子でエントリを限定することで、上述の入力限定方式と同様にフロー検索を実行することができる。

【0055】

【発明の効果】

【0056】本発明によると、以上のように、高通信品質(QoS)、高信頼性、セキュリティを保証して高速にルーティングするネットワーク中継装置及び方法を提供することができる。

【0057】また、本発明によると、ユーザを識別する情報やプロトコル情報や優先度情報等のフロー条件を大量に設定でき、回線速度の高速化やフロー条件の増加に対応することができ、高速にフロー検出して、高速にQoS制御やフィルタリングを実現することができる。また、フロー条件の記述性を向上させ、転送元及び転送先の情報等の組合せ条件により、優先制御、廃棄制御、帯域制御等の多種多様なフロー条件に柔軟に対応した制御を高速に処理することができる。さらに、本発明によると、質の異なる業務(基幹業務と情報系業務等)をひとつのネットワークに統合することができる。

【図面の簡単な説明】

【図1】本発明に係るネットワーク中継装置の構成図。

【図2】ルーティング処理部の内部構成を表したネットワーク中継装置の動作説明図。

【図3】ネットワーク中継装置の動作概要のシーケンス図。

【図4】パケットバッファ及びヘッダRAMの説明図。

【図5】経路検索に用いられる各テーブルの説明図。

【図6】ルーティング処理部による高速化処理の説明図。

【図7】ハードウェアによる検索エンジンの構成図。

【図8】パイプライン制御による高速処理の説明図。

【図9】フロー検索処理の説明図。

【図10】フロー検索テーブルの説明図。

【図11】第1の入力回線限定方式の説明図。

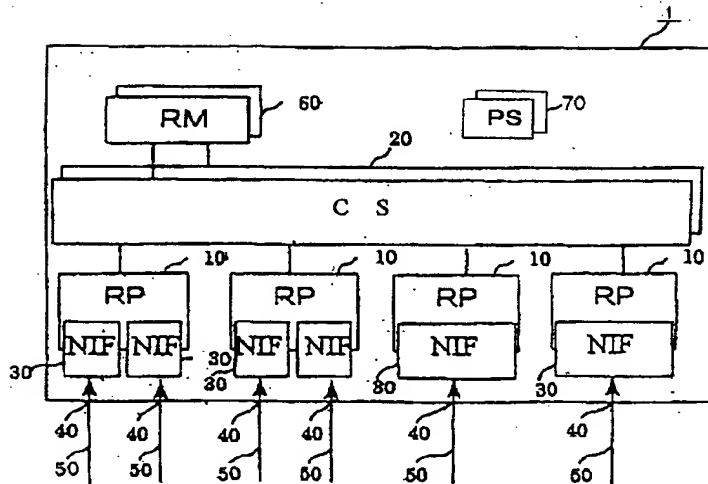
【図12】第2の入力回線限定方式の説明図。

【図13】従来のネットワーク中継装置の構成図。

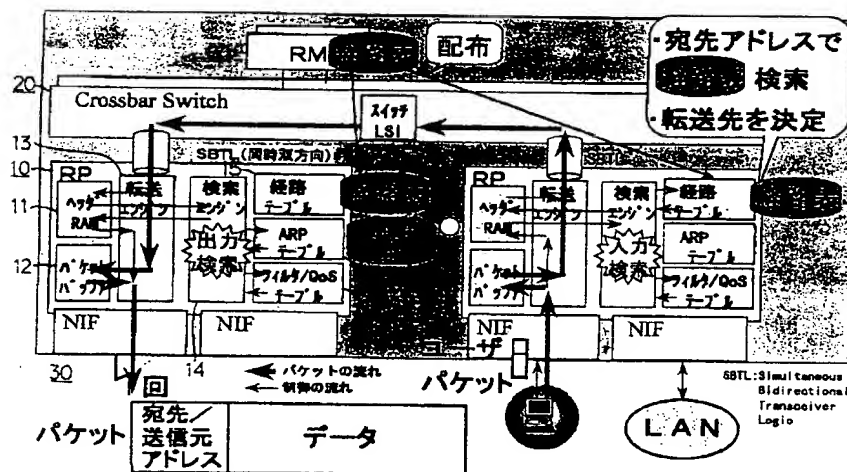
【符号の説明】

- 1 ネットワーク中継装置
- 10 ルーティング処理部
- 20 スイッチ部
- 30 ネットワークインタフェース部
- 40 ポート
- 50 ネットワーク
- 60 ルーティング管理部

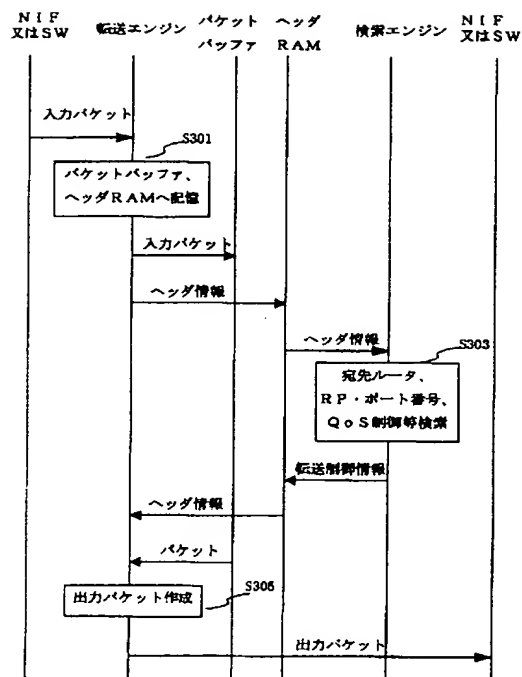
【図1】



【図2】

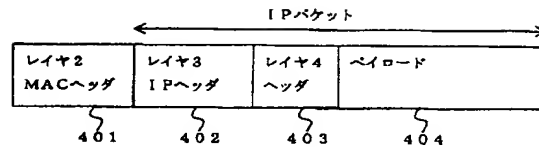


【図3】

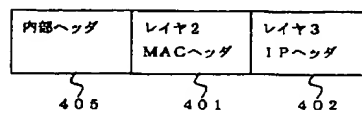


【図4】

(A) パケットのフォーマット

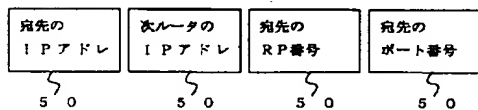


(B) ヘッダ情報のフォーマット

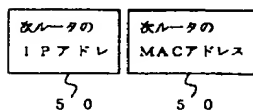


【図5】

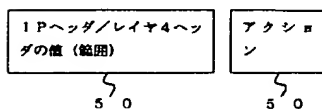
(A) 経路テーブルの各エントリのフォーマット



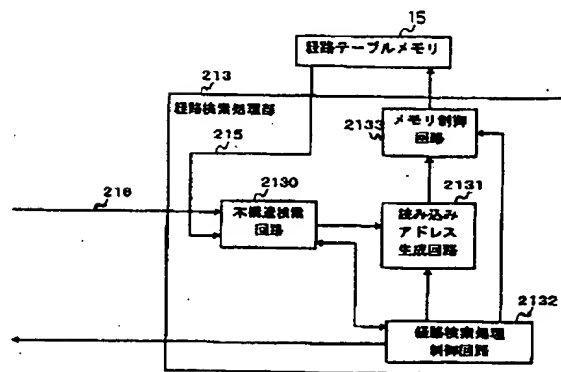
(B) ARPテーブルの各エントリのフォーマット



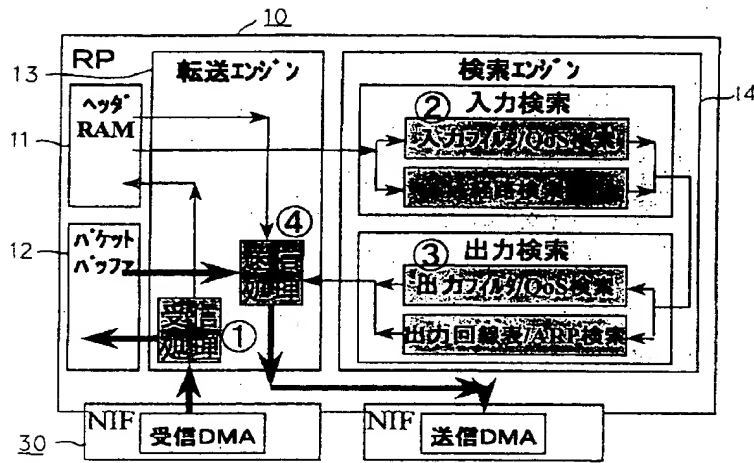
(C) フィルタ/QoSテーブルの各エントリのフォーマット



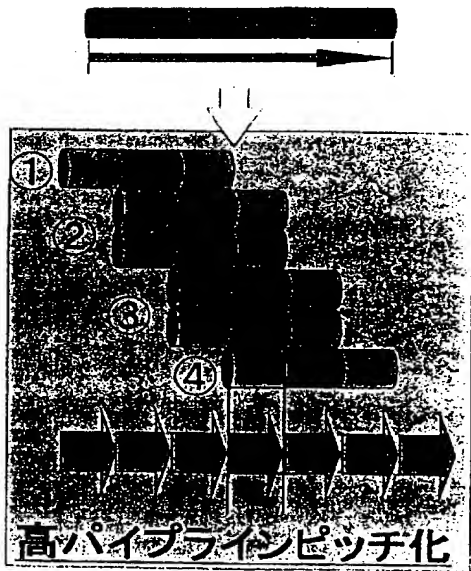
【図7】



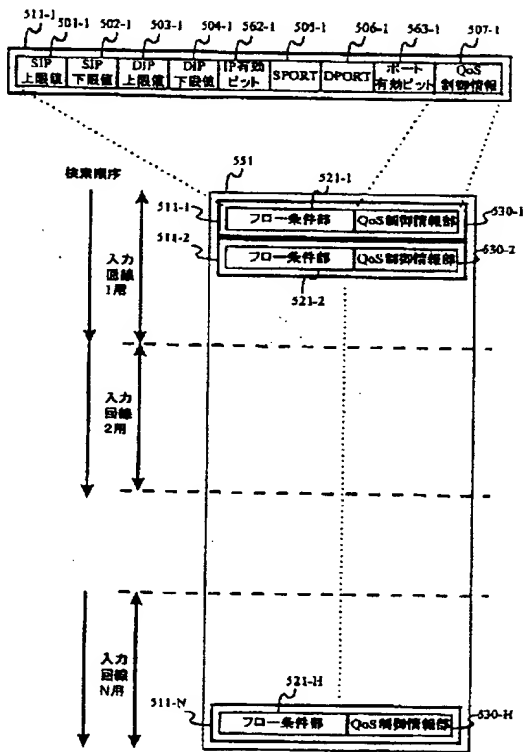
【図6】



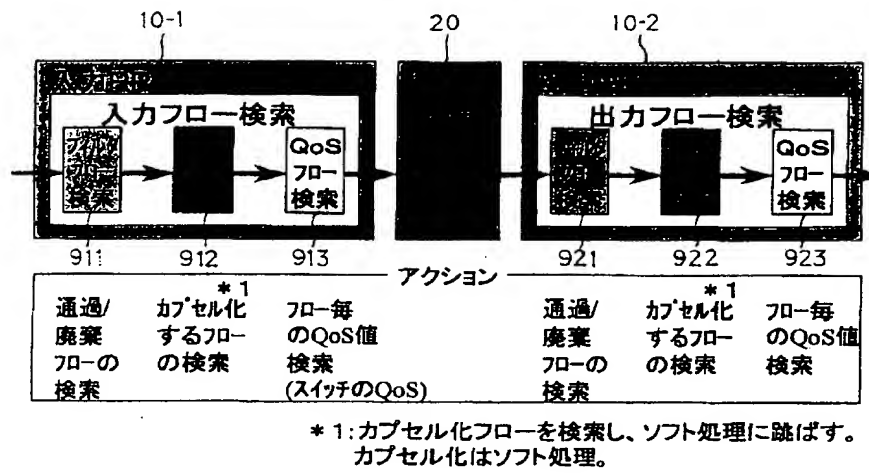
【図8】



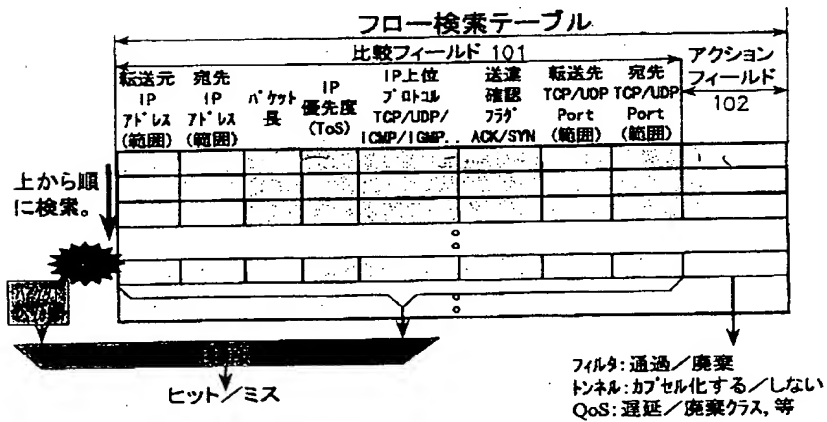
【図11】



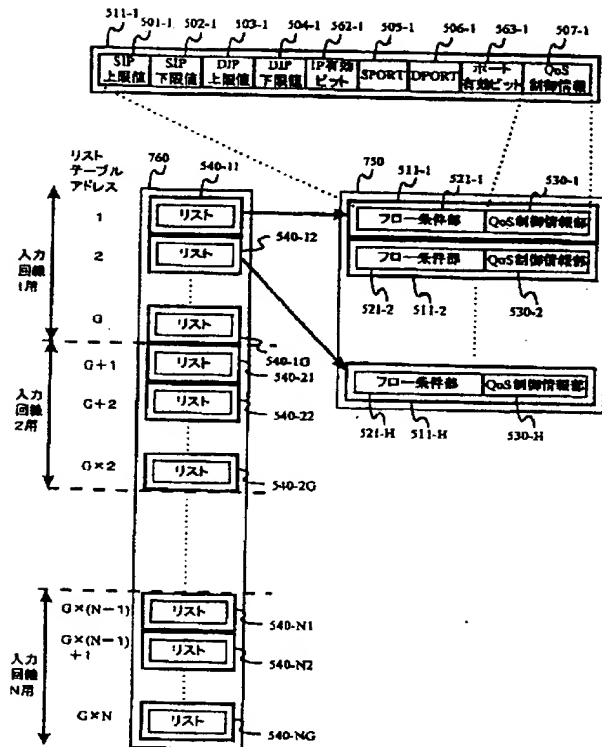
【図 9】



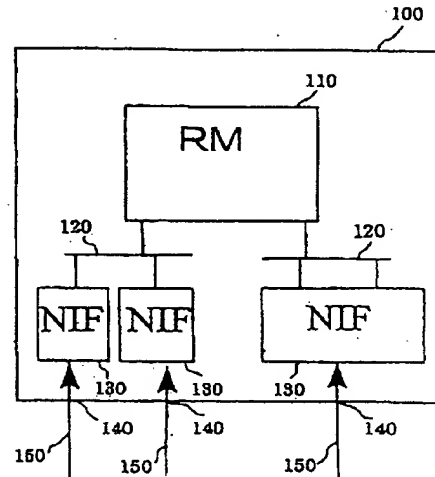
【図 10】



【図12】



【図13】



フロントページの続き

(72)発明者 相本 毅
東京都国分寺市東恋ヶ窪 1 丁目280番地
株式会社日立製作所中央研究所内

(72)発明者 矢崎 武己
東京都国分寺市東恋ヶ窪 1 丁目280番地
株式会社日立製作所中央研究所内

(72)発明者 松山 信仁
神奈川県秦野市堀山下 1 番地 株式会社日
立インフォメーションテクノロジー内

(72)発明者 左古 義人
神奈川県秦野市堀山下 1 番地 株式会社日
立製作所汎用コンピュータ事業部内

(72)発明者 田部 智彦
神奈川県秦野市堀山下 1 番地 株式会社日
立製作所汎用コンピュータ事業部内

F ターム(参考) 5K030 GA01 GA11 HA10 HB11 HB28
HC01 HC14 HD03 KA13 KX12
KX29 LB05 LC01 LC15 LE05
5K034 AA01 AA05 DD03 EE09 FF11
MM11 MM21
9A001 CC02 CC03 CC07 KK56 LL02
LL03